AFRL-RV-HA-TR-2008-1016

Error Analysis in the Joint Event Location/Seismic Calibration Inverse Problem

William L. Rodi

Earth Resources Laboratory
Department of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
77 Massachusetts Avenue, 54-514
Cambridge, MA 02139

Final Report

15 February 2008

AIR FORCE RESEARCH LABORATORY
Space Vehicles Directorate
29 Randolph Road
AIR FORCE MATERIEL COMMAND
Hanscom AFB, MA 01731-3010

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release and is available to the general public, including foreign nationals. Qualified requestors may obtain additional copies from the Defense Technical Information Center (DTIC) (http://www.dtic.mil). All others should apply to the National Technical Information Service.

AFRL-RV-HA-TR-2008-1016 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//Signed//	//Signed//
ROBERT RAISTRICK Contract Manager	PAUL TRACY, Acting Chief Battlespace Surveillance Innovation Center

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
15-02-2008	Scientific Report - Final	31-10-2003 to 30-09-2007

4. TITLE AND SUBTITLE
Error Analysis in the Joint Event Location/Seismic Calibration Inverse Problem

5a. CONTRACT NUMBER
F19628-03-C-0109

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
62601F

6. AUTHOR(S)
William L. Rodi

5d. PROJECT NUMBER
1010

5e. TASK NUMBER
SM

5f. WORK UNIT NUMBER
A1

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Earth Resources Laboratory
Dept. of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
77 Massachusetts Avenue, 54-514
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
29 Randolph Rd.
Hanscom AFB, MA 01731-3010

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RVBYE

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-RV-HA-TR-2008-1016

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
The goal of this project was to develop new mathematical and computational techniques for quantifying the errors in seismic event locations, focusing on the effects of errors in travel-time predictions from a velocity model. Our approach associates these model errors with the uncertainty in path travel-time corrections inferred from a calibration analysis. The main accomplishment of the project was a general formulation of location uncertainty in terms of the joint inverse problem that combines event location and seismic calibration. The formulation accommodates travel-time nonlinearity, Gaussian and non-Gaussian observational error, and a broad class of parameterizations of path travel-time corrections. We implemented the formulation for the simple correction parameterization as station time terms, applying it to data from the Nevada Test Site (NTS) as a proof of concept of the joint inversion approach. To address the computational intensity of the approach, we re-formulated it as an approximate, two-stage process - calibration followed by location - which we also implemented for the time-term case using newly designed algorithms for each stage. Application of the two-stage approach to NTS data demonstrated its much greater efficiency and suggest it as a feasible uncertainty paradigm for more complex problems in which travel-time corrections are parameterized with 3-D Earth models.

15. SUBJECT TERMS
Seismic event location, Seismic calibration, Location uncertainty, Confidence regions

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Robert J. Raistrick
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR	37	19b. TELEPHONE NUMBER *(include area code)* 781-377-3726

Contents

List of Figures

Summary

The goal of this project was to develop new mathematical and computational techniques for quantifying the errors in seismic event locations, with the goal of improving on the epicentral confidence ellipses and focal depth confidence intervals in conventional use. The primary focus of the project was to develop a rigorous approach to the effects of model errors, or errors in the travel-time predictions calculated from an assumed velocity model for the Earth. The approach that was investigated associates model errors with the uncertainty in path travel-time corrections that are inferred from a calibration analysis. This naturally leads to the idea of analyzing location uncertainty in the context of a joint inverse problem that treats event location and travel-time calibration as coupled problems.

A primary accomplishment of the project was a general theoretical formulation of the joint location/calibration inverse problem that applies to a wide range of situations, including the use of non-Gaussian distributions for pick errors, the accommodation of non-GT0 calibration events, and the use of a general parameterization of travel-time corrections that subsumes a variety of calibration techniques previous researchers have considered. The latter includes simple station-specific time terms, station-specific travel-time correction surfaces, and 3D velocity models. The formulation leads to a general expression for event location confidence regions based on the statistical framework of hypothesis testing with likelihood-ratio statistics.

A second accomplishment of the project was to implement the theoretical formulation for the case of time-term corrections as a proof of concept of the joint inversion uncertainty approach. Applications performed with data from Nevada Test Site explosions illustrate the effects of model errors on event location confidence regions and their dependence on the location accuracy of calibration events. Despite attempts to increase the computational efficiency of the joint inversion approach, these applications also indicated that the approach is not feasible for complex methods of travel-time calibration like 3-D velocity tomography.

The third accomplishment of the project was the development of an approximate, but much more efficient, version of the new uncertainty approach that follows the traditional factorization of travel-time calibration and event location into a two-stage process. The approximation assumes that the uncertainty in travel-time corrections estimated in a calibration analysis can be represented by a Gaussian probability distribution. The generality of the theoretical formulation is otherwise preserved. To test the two-stage approach, new algorithms were developed for calculating calibration uncertainty (stage one) and employing it in the calculation of location confidence regions (stage two). Application of these algorithms to the Nevada Test Site data, assuming time-term corrections, demonstrate the much greater efficiency of the two-stage approach and suggest that the approach is feasible for the case of tomographic calibration.

Chapter 1

Introduction

The assessment of errors in seismic event locations remains a challenging problem for nuclear monitoring research. Theoretical treatments of location uncertainly have generally focused on the effects of observational, or pick, errors in the seismic arrival times (or other observed attributes) used in locating an event. However, it has been recognized for a long time that an equal or larger source of location uncertainty, especially when arrivals at regional distances are involved, is the error in the travel-time predictions that are calculated with models of the Earth's velocity structure. Formulations of location uncertainty have generally lacked a rigorous treatment of these model errors, using instead the ad hoc assumption that they simply augment the pick errors.

The project reported here has attempted to develop a rigorous, and general, treatment of model errors by linking them to the errors incurred in the calibration of seismic travel-times. This association leads naturally to consideration of a joint inverse problem in which arrival time data from multiple events are used to infer the event locations together with calibration parameters that generate corrections to the model-based travel-time predictions. Taking one of the multiple events to be the target of investigation, and the others as historical calibration events, an error analysis performed on the joint location/calibration inverse problem implicitly accounts for the effects of model errors on the target event location. The joint inversion approach also provides a means to address the important issue of how errors in the locations of calibration events influence the location uncertainty for the target event.

The next chapter of this report formulates the problem of event location uncertainty and reviews the classical theory of elliptical confidence regions on locations, which follows from the linearization of the travel-time forward problem and the assumption that pick errors obey a Gaussian probability distribution. The linear/Gaussian theory is then generalized based on the statistical framework of maximum-likelihood estimation and hypothesis testing, and numerical techniques for the evaluation of non-elliptical confidence regions are described.

Chapter 3 of the report extends the theory and algorithms from single-event location to multiple-event location, i.e. the joint location/calibration problem. The resulting approach is illustrated with examples that use data from well-located explosions at the Nevada Test Site.

Chapter 4 presents a two-stage method that, following conventional practice, separates the calibration and location aspects of the joint inverse problem. The separation is developed under a specific approximation to the full theory; namely, that the uncertainty in calibration parameters can be represented by a multivariate Gaussian distribution. Appropriate algorithms to implement the two-stage approach are presented and illustrated with the Nevada Test Site data. The report ends with some conclusions and suggestions for future work.

Chapter 2

Uncertainty Analysis in Single-Event Location

2.1 The Single-Event Location Problem and Approaches to Uncertainty

Most event location algorithms infer the hypocenter and origin time of an event from a set of arrival times of various seismic phases observed at various stations of a network. We will denote the arrival-time data d_i for $i = 1, \ldots, n$, where n is the total number of station/phase combinations that have been observed. Letting $\mathbf{x}$ denote the event hypocenter and t its origin time, we can express the single-event location problem as

$$d_i = T_i(\mathbf{x}) + t + e_i, \quad i = 1, \ldots, n, \tag{2.1}$$

where e_i is the observational error in d_i, and T_i is a function that predicts the travel-time of the ith station/phase pair as a function of the event hypocenter. The travel-time prediction functions are usually based on a velocity model for the Earth. In matrix/vector notation we can write equation (2.1) as

$$\mathbf{d} = \mathbf{T}(\mathbf{x}) + \mathbf{1}t + \mathbf{e}, \tag{2.2}$$

where each element of the column vector $\mathbf{1}$ is unity.

A solution of the single-event location problem is commonly taken to comprise point estimates of the event location parameters, which we will denote $(\widehat{\mathbf{x}}, \widehat{t})$, together with a probabilistic description of the uncertainty in these estimates. The uncertainty description follows from an assumed probabilistic model for the observational errors, e_i. There are a number of accepted ways to describe uncertainty. For example, if $\widehat{\mathbf{x}}$ is found to be a Gaussian random variable, its uncertainty is well-characterized by a variance-covariance matrix, $\mathbf{V}$, such that

$$\mathrm{E}[(\widehat{\mathbf{x}} - \mathbf{x})(\widehat{\mathbf{x}} - \mathbf{x})^T] = \mathbf{V}, \tag{2.3}$$

where $\mathrm{E}[\ ]$ denotes the mathematical expectation of a random variable. The presumption here is that $\widehat{\mathbf{x}}$ is an *unbiased* estimate of $\mathbf{x}$, or $\mathrm{E}[\widehat{\mathbf{x}}] = \mathbf{x}$. There is no guarantee, however, that an unbiased Gaussian estimate of $\mathbf{x}$ exists or, if it does, that its variance is independent of $\mathbf{x}$.

A more general description of uncertainty is accomplished with Neyman-Pearson confidence regions on the unknown parameters or subsets thereof. A confidence region on $\mathbf{x}$, for example, is a region of $\mathbf{x}$-space, calculated from the observed data, to which we can assign a probability of

including the true hypocenter $\mathbf{x}$. Letting β be that probability, or *confidence level* (e.g. $\beta = 0.9$ for a 90% confidence region), and letting $\mathcal{X}_\beta(\mathbf{d})$ be its associated confidence region, we have

$$\text{Prob}\left[\mathbf{x} \in \mathcal{X}_\beta(\mathbf{d})\right] = \beta. \tag{2.4}$$

The probability operator used here is that which is induced by the assumed error model for $\mathbf{e}$ and, hence, $\mathbf{d}$.

Bayesian inference provides another approach to uncertainty. In this approach, the location parameters are considered to be random variables, and a complete solution of the inverse problem is embodied in their joint *posterior* probability density function (p.d.f.), denoted $f[\mathbf{x}, t \mid \mathbf{d}]$ (where "$|\mathbf{d}$" indicates conditioning on the data vector). From the posterior distribution, a point estimate of $\mathbf{x}$ is available as the MAP (maximum a posteriori) estimate:

$$f[\widehat{\mathbf{x}} \mid \mathbf{d}] = \max_{\mathbf{x}} f[\mathbf{x} \mid \mathbf{d}], \tag{2.5}$$

where $f[\mathbf{x} \mid \mathbf{d}]$ is the *marginal* posterior p.d.f. of $\mathbf{x}$, defined as

$$f[\mathbf{x} \mid \mathbf{d}] = \int dt \; f[\mathbf{x}, t \mid \mathbf{d}]. \tag{2.6}$$

A Bayesian confidence region can be defined from the posterior p.d.f. with

$$\int_{\mathbf{x} \in \mathcal{X}_\beta(\mathbf{d})} d\mathbf{x} \; f[\mathbf{x} \mid \mathbf{d}] = \beta \tag{2.7}$$

but such confidence regions do not necessarily obey the *confidence property* stated in equation (2.4).

This project has focused on Neyman-Pearson confidence regions as a framework for event location uncertainty analysis. Before describing this approach in a general context, we describe it for the special case most often assumed in seismic event location studies.

2.2 Gaussian/Linear Confidence Regions

The traditional solution to the single-event location problem consists of *least-squares* estimates of the event parameters. These minimize a data-misfit function given by

$$\Psi(\mathbf{x}, t; \mathbf{d}) = \sum_{i=1}^{n} w_i \; \left(d_i - T_i(\mathbf{x}) - t\right)^2, \tag{2.8}$$

where the w_i are prescribed weights. We can also write the least-squares misfit function as

$$\Psi(\mathbf{x}, t; \mathbf{d}) = \left(\mathbf{d} - \mathbf{T}(\mathbf{x}) - \mathbf{1}t\right)^T \mathbf{W} \left(\mathbf{d} - \mathbf{T}(\mathbf{x}) - \mathbf{1}t\right) \tag{2.9}$$

with the diagonal matrix $\mathbf{W}$ containing the weights:

$$\mathbf{W} = \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_n \end{pmatrix}. \tag{2.10}$$

The least-squares parameter estimates satisfy

$$\Psi(\widehat{\mathbf{x}}, \widehat{t}; \mathbf{d}) = \min_{\mathbf{x}, t} \Psi(\mathbf{x}, t; \mathbf{d}). \tag{2.11}$$

Numerically, the minimization of Ψ is typically performed with a Gauss-Newton algorithm, as proposed originally by Geiger (1912) and re-formulated by Flinn (1965).

Flinn (1965) also developed the theory for location confidence regions in the least-squares setting, and we redevelop this result here as a prelude to a more general theory. Flinn derives Neyman-Pearson confidence regions in a hypothesis testing framework. The gist of this approach is as follows. Given an arbitrary hypocenter $\mathbf{x}$, we take as a null hypothesis that the true event hypocenter is $\mathbf{x}$. If, based on a test statistic $\tau(\mathbf{x}, \mathbf{d})$, we *cannot reject* the null hypothesis at some level of confidence, we take $\mathbf{x}$ to be a point *inside* the confidence region at that level.

Flinn applied this approach with the underlying assumption that the observational errors are zero-mean, Gaussian random variables whose standard deviations are known within a constant factor, σ. The probability density function (p.d.f.) of e_i in this case can be written as

$$f[e_i] = \frac{1}{(2\pi)^{1/2}\sigma\nu_i} \exp\left\{-\frac{1}{2}\left(\frac{e_i}{\sigma\nu_i}\right)^2\right\}, \tag{2.12}$$

where ν_i is a nominal standard deviation assigned to e_i. With this assumption, and assuming that the errors are statistically independent (uncorrelated) the weights for the data-misfit function are naturally set as

$$w_i \sim \nu_i^{-2}. \tag{2.13}$$

Location confidence regions can be defined on all the event location parameters ($\mathbf{x}$ and t) in four-dimensional space, or on any subset of the parameters. Here we will consider confidence regions on the hypocenter $\mathbf{x}$, thus treating t as a "nuisance" parameter. For these, Flinn (1965) applied hypothesis testing with the test statistic given by

$$\tau(\mathbf{x}, \mathbf{d}) = \frac{1}{\widehat{\sigma}^2}\left[\min_t \Psi(\mathbf{x}, t; \mathbf{d}) - \Psi(\widehat{\mathbf{x}}, \widehat{t}; \mathbf{d})\right], \tag{2.14}$$

where $(\widehat{\mathbf{x}}, \widehat{t})$ is the least-squares location solution and $\widehat{\sigma}$ is an estimate of σ which Flinn took to be

$$\widehat{\sigma}^2 = \frac{1}{n-4}\Psi(\widehat{\mathbf{x}}, \widehat{t}; \mathbf{d}). \tag{2.15}$$

We see that τ compares the data misfit achieved by a given hypocenter $\mathbf{x}$ (allowing t to adjust) to the minimum misfit achieved by any $\mathbf{x}$. A confidence region on $\mathbf{x}$ at the β confidence level, or $\mathcal{X}_\beta(\mathbf{d})$, comprises those points $\mathbf{x}$ yielding values of the test statistic below some cutoff. That is,

$$\mathcal{X}_\beta(\mathbf{d}) = \{\mathbf{x}\colon \tau(\mathbf{x}, \mathbf{d}) \le \tau_\beta\}, \tag{2.16}$$

where τ_β is a critical value of the test statistic satisfying

$$\text{Prob}\left[\tau(\mathbf{x}, \mathbf{d}) \le \tau_\beta\right] = \beta. \tag{2.17}$$

Equation (2.16) can be interpreted to mean that $\mathbf{x}$ can be rejected as the true hypocenter if its relative data misfit, as measured by $\tau(\mathbf{x}, \mathbf{d})$, is too large. Equation (2.17) states that the cutoff value for this rejection, τ_β, marks the β point on the cumulative probability distribution of $\tau(\mathbf{x}, \mathbf{d})$, as determined under the null hypothesis. This choice for the cutoff ensures the confidence property of equation (2.4). That is, the true hypocenter will be rejected with probability $1 - \beta$ or, stated conversely, the confidence region will include the true hypocenter with

probability β. This interpretation presumes, however, that the distribution of $\tau(\mathbf{x}, \mathbf{d})$ does not depend on the parameters t and σ, which we will see is indeed the case in the Gaussian/linear analysis.

To obtain the confidence region on $\mathbf{x}$ in a convenient form we need a numerical value for τ_β *and* a way to parameterize the points satisfying (2.16). These requirements are easily met under the Gaussian error assumption if we make the additional assumption that the travel-time functions can be well-approximated by a linear expansion about the least-squares estimate:

$$\mathbf{T}(\mathbf{x}) = \mathbf{T}(\widehat{\mathbf{x}}) + \mathbf{A}(\mathbf{x} - \widehat{\mathbf{x}}). \tag{2.18}$$

Here, $\mathbf{A}$ is a Jacobian matrix containing derivatives of the travel-time functions, evaluated at $\mathbf{x} = \widehat{\mathbf{x}}$:

$$A_{ij} = \left.\frac{\partial T_i(\mathbf{x})}{\partial x_j}\right|_{\mathbf{x}=\hat{\mathbf{x}}}. \tag{2.19}$$

Given (2.18), the location problem becomes a linear inverse problem (the dependence of t is exactly linear), and standard methods of Gaussian/linear analysis can be applied to determine τ and its probability distribution (e.g., Anderson, 1965). We report the results.

Under the linearization of $\mathbf{T}$, the test statistic can be expressed as

$$\tau(\mathbf{x}, \mathbf{d}) = \frac{1}{\widehat{\sigma}^2}(\mathbf{x} - \widehat{\mathbf{x}})^T \mathbf{A}^T \mathbf{Q}^T \mathbf{W} \mathbf{Q} \mathbf{A} (\mathbf{x} - \widehat{\mathbf{x}}), \tag{2.20}$$

where $\mathbf{Q}$ is the projection matrix given by

$$\mathbf{Q} = \mathbf{I} - \frac{1}{\mathbf{1}^T \mathbf{W} \mathbf{1}} \mathbf{1}\mathbf{1}^T \mathbf{W}. \tag{2.21}$$

Given that $\mathbf{e}$ is Gaussian, τ is the (scaled) ratio of two independent chi-squared random variables having 3 and $n-4$ degrees of freedom, respectively. τ itself (divided by 3) is thus F distributed with these degrees of freedom, and we find that

$$\tau_\beta = 3\, F_\beta(3, n-4). \tag{2.22}$$

$F_\beta(k, m)$ denotes the β-point of the F distribution for k and m degrees of freedom. The confidence region on $\mathbf{x}$ at confidence level β is thus given by

$$(\mathbf{x} - \widehat{\mathbf{x}})^T \mathbf{A}^T \mathbf{Q}^T \mathbf{W} \mathbf{Q} \mathbf{A} (\mathbf{x} - \widehat{\mathbf{x}}) \le 3\, \widehat{\sigma}^2\, F_\beta(3, n-4). \tag{2.23}$$

We see from equation (2.22) that τ_β does not depend on any of the unknown parameters $(\mathbf{x}, t, \sigma)$. Equation (2.23) therefore describes the interior of an ellipsoid in hypocenter space, centered at the least-squares estimate $\widehat{\mathbf{x}}$. The parameters of the ellipsoid can be determined from the eigenvectors and eigenvalues of the matrix $\mathbf{A}^T \mathbf{Q}^T \mathbf{W} \mathbf{Q} \mathbf{A}$. The inverse square-root of the eigenvalues, multiplied by the square-root of the right-hand side of (2.23), are the semi-axis lengths of the ellipsoid.

An alternative to Flinn's (1965) results was proposed by Evernden (1969) for the case in which σ is taken to be a known quantity. In this case, $\tau(\mathbf{x}, \mathbf{d})$ in equation (2.20), with $\widehat{\sigma}$ set to the known σ, becomes chi-squared distributed with 3 degrees of freedom and the confidence region becomes

$$(\mathbf{x} - \widehat{\mathbf{x}})^T \mathbf{A}^T \mathbf{Q}^T \mathbf{W} \mathbf{Q} \mathbf{A} (\mathbf{x} - \widehat{\mathbf{x}}) \le \sigma^2\, \chi^2_\beta(3). \tag{2.24}$$

The confidence region scaling is smaller when the chi-squared distribution is used for τ, signifying that a loss of information occurs when σ is not known and has to be estimated from the data. Table 2.1 shows the ratio of confidence region scaling factors implied by Flinn's and Evernden's methods.

Table 2.1: Ratio of F to χ^2 scaling of hypocenter confidence regions

$$\text{Ratio} = \left[\frac{3F_\beta(3, n-4)}{\chi^2_\beta(3)}\right]^{1/2}$$

n	$\beta = 90\%$	$\beta = 95\%$
5	5.071	9.100
6	2.097	2.712
8	1.418	1.591
10	1.256	1.351
12	1.185	1.249
15	1.130	1.174
20	1.087	1.115
30	1.052	1.069
50	1.029	1.038

2.3 Non-Elliptical Confidence Regions

In previous work (e.g. Rodi and Toksöz, 2000), we generalized the linear/Gaussian theory for location confidence regions to allow non-Gaussian probability distributions for pick errors, to accommodate nonlinear constraints on location parameters, and to take account of the nonlinearity of the travel-time functions. We summarize this generalization here.

2.3.1 Likelihood function for arrival data

We still assume the pick errors are statistically independent but, following Billings *et al.* (1994), allow each to be distributed with a *generalized Gaussian* distribution, whose probability density function (p.d.f.) is given by

$$f[e_i] = \frac{1}{K(p)\sigma_i} \exp\left\{-\frac{1}{p}\left|\frac{e_i}{\sigma_i}\right|^p\right\}, \quad p \geq 1, \tag{2.25}$$

where σ_i is a standard error and

$$K(p) = 2p^{1/p}\Gamma(1 + 1/p). \tag{2.26}$$

(Γ is the gamma function.) When $p = 2$, the p.d.f. is Gaussian (normal) with a mean of zero and variance of σ_i^2. When $p = 1$, $f[\ ]$ is a Laplace distribution (two-sided exponential). Like we did earlier, we assume that the standard errors are known within a scale parameter σ:

$$\sigma_i = \sigma\nu_i \tag{2.27}$$

with the ν_i being known, nominal standard errors.

Given that the errors e_i are independent, the joint p.d.f. of the observed data d_i is determined by the product of the $f[e_i]$, i.e.

$$f[d_1, \ldots, d_n; \mathbf{x}, t, \sigma] = \prod_{i=1}^{n} f[e_i = d_i - T_i(\mathbf{x}) - t]. \tag{2.28}$$

Considered as a function of the unknown problem parameters, the data p.d.f. serves as a *likelihood* function. Equation (2.25) implies that the negative logarithm of the likelihood function is

$$\Lambda(\mathbf{x}, t, \sigma; \mathbf{d}) = n \log K(p) + \sum_{i=1}^{n} \log \nu_i + n \log \sigma + \frac{1}{p\sigma^p} \Psi(\mathbf{x}, t; \mathbf{d}) \tag{2.29}$$

where the data-misfit function now generalizes equation (2.8) as

$$\Psi(\mathbf{x}, t; \mathbf{d}) = \sum_{i=1}^{n} |d_i - T_i(\mathbf{x}) - t|^p / \nu_i^p. \tag{2.30}$$

We will often refer to the function Λ simply as the *penalty* function to avoid the more clumsy *negative log-likelihood*.

The generalization of the least-squares solution to the location problem is the *maximum-likelihood* solution, which minimizes the penalty function:

$$\Lambda(\widehat{\mathbf{x}}, \widehat{t}, \widehat{\sigma}; \mathbf{d}) = \min_{\mathbf{x}, t, \sigma} \Lambda(\mathbf{x}, t, \sigma; \mathbf{d}). \tag{2.31}$$

We have included $\widehat{\sigma}$, the maximum-likelihood estimate of σ, as part of the solution, replacing the unbiased estimate considered by Flinn (1965).

In practice, the minimization in equation (2.31) is performed over a restricted range of parameter values. Doing so effects hard parameter constraints which can represent important prior information, e.g. preventing event locations above the Earth's surface or below a maximum credible depth. Additionally, it is usually credible to constrain σ based on knowledge of the data processing procedures for picking arrival times. In the rest of this report, it will be understood that the minimization of a penalty function is subject to appropriate *a priori* constraints on the free parameters.

2.3.2 Neyman-Pearson confidence regions

Our more general formulation follows the Gaussian/linear uncertainty analysis described above and defines confidence regions in terms of hypothesis testing, as per equations (2.16) and (2.17). We now define the test statistic more generally as the logarithm of a likelihood ratio. For a confidence region on $\mathbf{x}$ we have

$$\tau(\mathbf{x}, \mathbf{d}) = \min_{t, \sigma} \Lambda(\mathbf{x}, t, \sigma; \mathbf{d}) - \Lambda(\widehat{\mathbf{x}}, \widehat{t}, \widehat{\sigma}; \mathbf{d}), \tag{2.32}$$

where Λ is the negative log-likelihood function of equation (2.29). We can also write this as

$$\tau(\mathbf{x}, \mathbf{d}) = \widetilde{\Lambda}(\mathbf{x}; \mathbf{d}) - \widetilde{\Lambda}(\widehat{\mathbf{x}}; \mathbf{d}), \tag{2.33}$$

where $\widetilde{\Lambda}$ is a "reduced" penalty function that is, for any fixed $\mathbf{x}$, minimized with respect to t and σ:

$$\widetilde{\Lambda}(\mathbf{x}; \mathbf{d}) = \min_{t, \sigma} \Lambda(\mathbf{x}, t, \sigma; \mathbf{d}). \tag{2.34}$$

We note that this statistic, in the Gaussian case ($p = 2$), does not default to the statistic used by Flinn (1965) (equation (2.14)). However, the statistics are equivalent in this situation and lead to the same confidence region for any β.

When the pick errors are not Gaussian ($p \neq 2$), and accepting the travel-time functions T_i as nonlinear, we cannot dismiss the possibility that the distribution of $\tau(\mathbf{x}, \mathbf{d})$ depends on some or all of the problem unknowns. Therefore, we now allow τ_β to depend on the unknown parameters and replace equation (2.16) with

$$\mathcal{X}_\beta(\mathbf{d}) = \{\mathbf{x}\colon \tau(\mathbf{x}, \mathbf{d}) \leq \widetilde{\tau}_\beta(\mathbf{x})\}, \tag{2.35}$$

where

$$\widetilde{\tau}_\beta(\mathbf{x}) = \max_{t,\sigma} \tau_\beta(\mathbf{x}, t, \sigma). \tag{2.36}$$

This differs from equation (2.16) in two regards. First, the dependence of τ_β on $\mathbf{x}$ lends a more complex nature to the inequality that delimits a confidence region. Second, the dependence on t and σ implies that the null hypothesis being tested (that $\mathbf{x}$ is the true hypocenter) depends on t and σ as nuisance parameters. Equations (2.35) and (2.36) define a *worst case* confidence region for this situation, i.e., the region is the largest that any values for the nuisance parameters can yield. A consequence of this definition is that the confidence property of equation (2.4) is not guaranteed to hold. However, a weaker version of the property does hold:

$$\text{Prob}\left[\mathbf{x} \in \mathcal{X}_\beta(\mathbf{d})\right] \geq \beta. \tag{2.37}$$

In this case one can state that the confidence region $\mathcal{X}_\beta(\mathbf{d})$ contains the true hypocenter with a probability of *at least* β.

2.3.3 Confidence regions on epicenter and focal depth

While we have thus far presented our methodology with respect to *hypocentral* confidence regions, and will continue to do so, we pause to show how the methodology applies to confidence regions on other event parameters. To do this, we separate the problem parameters into two types. The first are *target* parameters whose uncertainty we wish to characterize. The remaining parameters are *nuisance* parameters, which are not of interest but whose effect on the data cannot be ignored. Denoting the target parameters as components of a vector $\mathbf{p}$, and the nuisance parameters as components of $\mathbf{q}$, our discussion so far has focused on the situation

$$\mathbf{p} = \mathbf{x} \tag{2.38}$$
$$\mathbf{q} = (t, \sigma). \tag{2.39}$$

Denoting $\mathbf{x} = (\theta, \phi, z)$, where θ, ϕ and z are latitude, longitude and depth, respectively, we can consider confidence regions on the event epicenter (θ, ϕ) by letting

$$\mathbf{p} = (\theta, \phi) \tag{2.40}$$
$$\mathbf{q} = (z, t, \sigma), \tag{2.41}$$

and on the event focal depth by letting

$$\mathbf{p} = z \tag{2.42}$$
$$\mathbf{q} = (\theta, \phi, t, \sigma). \tag{2.43}$$

Our uncertainty formulation transforms to $\mathbf{p}$ and $\mathbf{q}$ as follows. First, we re-denote the penalty function as $\Lambda(\mathbf{p}, \mathbf{q}; \mathbf{d})$ and the test statistic for hypothesis testing as $\tau(\mathbf{p}, \mathbf{d})$. We have

$$\tau(\mathbf{p}, \mathbf{d}) = \widetilde{\Lambda}(\mathbf{p}; \mathbf{d}) - \widetilde{\Lambda}(\widehat{\mathbf{p}}; \mathbf{d}), \tag{2.44}$$

where $(\widehat{\mathbf{p}}, \widehat{\mathbf{q}})$ is the maximum-likelihood solution and

$$\widetilde{\Lambda}(\mathbf{p}; \mathbf{d}) = \min_{\mathbf{p}} \Lambda(\mathbf{p}, \mathbf{q}; \mathbf{d}). \tag{2.45}$$

The confidence region on $\mathbf{p}$ at confidence level β is then given by

$$\mathcal{P}_\beta(\mathbf{d}) = \{\mathbf{p} \colon \tau(\mathbf{p}, \mathbf{d}) \leq \widetilde{\tau}_\beta(\mathbf{p})\} \tag{2.46}$$

with

$$\widetilde{\tau}_\beta(\mathbf{p}) = \max_{\mathbf{q}} \tau_\beta(\mathbf{p}, \mathbf{q}). \tag{2.47}$$

We will resume the description of the methodology for hypocentral confidence regions, corresponding to the case of equations (2.38) and (2.39).

2.4 Numerical Algorithm for Confidence Regions

Our general assumptions do not permit an analytic solution for $\widetilde{\tau}_\beta$ in equation (2.36) or a simple geometric description of the confidence region given by equation (2.35). To work around this, Rodi and Toksöz (2000) devised a numerical approach to computing and representing event location confidence regions, which we now describe.

The algorithm involves two steps: (1) mapping $\tau(\mathbf{x}, \mathbf{d})$ on a grid in $\mathbf{x}$-space, and (2) for given confidence level β, estimating $\widetilde{\tau}_\beta(\mathbf{x})$ for each point on the grid. After performing these two steps, the confidence region at the β level is represented by flagging the grid points for which $\tau(\mathbf{x}, \mathbf{d}) \leq \widetilde{\tau}_\beta(\mathbf{x})$.

The first step, which we will call *likelihood mapping*, entails minimizing the penalty function with respect to t and σ with $\mathbf{x}$ fixed, repeating for each $\mathbf{x}$ on the hypocenter grid.

In this project we considered two ways of performing the second step, which we now discuss.

2.4.1 Monte Carlo simulation

For any assumed value of σ (within prescribed bounds, if any), one can generate pseudo-random realizations of the error vector $\mathbf{e}$ based on the probability distribution for errors prescribed by equations (2.25) and (2.27). We will denote such a realization $\mathbf{e}^{\mathsf{mc}}(\sigma)$. Then, for given $\mathbf{x}$ and t, a realization of data is given by

$$\mathbf{d}^{\mathsf{mc}} = \mathbf{T}(\mathbf{x}) + \mathbf{1}t + \mathbf{e}^{\mathsf{mc}}(\sigma). \tag{2.48}$$

A realization of τ is obtained from this data as

$$\tau^{\mathsf{mc}}(\mathbf{x}, \mathbf{d}^{\mathsf{mc}}) = \widetilde{\Lambda}(\mathbf{x}; \mathbf{d}^{\mathsf{mc}}) - \min_{\mathbf{x}} \widetilde{\Lambda}(\mathbf{x}; \mathbf{d}^{\mathsf{mc}}). \tag{2.49}$$

We see that the computation of τ^{mc} entails performing event location (minimization of the penalty function) using $\mathbf{d}^{\mathsf{mc}}$ as synthetic data. Generating many realizations of τ in this way, one can use their histogram to estimate the critical statistic value, $\tau_\beta(\mathbf{x}, t, \sigma)$, for particular values of β. We have found that on the order of 300 realizations of τ^{mc} provide adequate accuracy for τ_β.

Strict adherence to equations (2.35) and (2.36), however, requires performing a Monte Carlo simulation, not just for each grid point $\mathbf{x}$, but for a sufficient number of t and σ values to be able to find a maximum of τ_β over these variables. The Monte Carlo approach thus becomes very

computationally intensive. We have therefore implemented this approach with the major simplification that the simulation be performed with only *one* combination of the parameters. The parameter values used are the maximum-likelihood estimates. Thus, synthetic data realizations are generated as

$$\mathbf{d}^{\mathsf{mc}} = \mathbf{T}(\widehat{\mathbf{x}}) + \mathbf{1}\widehat{t} + \mathbf{e}^{\mathsf{mc}}(\widehat{\sigma}), \tag{2.50}$$

and the formula for a confidence region becomes

$$\mathcal{X}_\beta(\mathbf{d}) = \left\{\mathbf{x}\colon \tau(\mathbf{x},\mathbf{d}) \le \tau_\beta(\widehat{\mathbf{x}},\widehat{t},\widehat{\sigma})\right\}. \tag{2.51}$$

We point out that other investigators, notably Wilcock and Toomey (1991), have used the same simplification in similar applications.

Ignoring the dependence of τ_β on origin time is quite justified since this dependence arises only if bounds on t are invoked, which is not normally done. The dependence on $\mathbf{x}$ and σ is not so easily dismissed. We can infer from the Gaussian/linear analysis that the dependence on $\mathbf{x}$ and σ will be significant only if the nonlinearity of the travel-time functions, T_i, is significant or when the generalize Gaussian order p differs significantly from 2.

2.4.2 Likelihood integration

The Bayesian approach to uncertainty suggests an alternative approximation to $\widetilde{\tau}_\beta(\mathbf{x})$. Let us assume a vacuous prior distribution on the unknown parameters, or $f[\mathbf{x},t,\sigma] \propto 1$. The likelihood function then becomes an unnormalized posterior p.d.f. on the parameters, or

$$f[\mathbf{x},t,\sigma \mid \mathbf{d}] = K(\mathbf{d})\, e^{-\Lambda(\mathbf{x},t,\sigma;\mathbf{d})}, \tag{2.52}$$

where $K(\mathbf{d})$ is a normalizing factor. A Bayesian confidence region on $\mathbf{x}$ (or other target parameter) is determined by the marginal posterior p.d.f. of $\mathbf{x}$, as we indicated earlier in equation (2.7). Let us instead consider a *quasi*-Bayesian approach in which the marginal posterior of $\mathbf{x}$ is replaced with a distribution that is *maximized* with respect to nuisance parameters. Define

$$\widetilde{f}[\mathbf{x} \mid \mathbf{d}] = \max_{t,\sigma} f[\mathbf{x},t,\sigma \mid \mathbf{d}]. \tag{2.53}$$

A quasi-Bayesian confidence region on $\mathbf{x}$ satisfies, in analogy with (2.7),

$$\int_{\mathbf{x}\in\mathcal{X}_\beta(\mathbf{d})} d\mathbf{x}\ \widetilde{f}[\mathbf{x} \mid \mathbf{d}] = \beta. \tag{2.54}$$

Given (2.33) and (2.34), it is not hard to show that

$$\widetilde{f}[\mathbf{x} \mid \mathbf{d}] = K(\mathbf{d})\, e^{-\widetilde{\Lambda}(\mathbf{x};\mathbf{d})} = K'(\mathbf{d})\, e^{-\tau(\mathbf{x};\mathbf{d})}, \tag{2.55}$$

where $K'(\mathbf{d})$ is a different normalizing factor. If we require that $\mathcal{X}_\beta(\mathbf{x})$ discriminate on the p.d.f. value—i.e.

$$\mathcal{X}_\beta(\mathbf{d}) = \left\{\mathbf{x}\colon \widetilde{f}[\mathbf{x} \mid \mathbf{d}] \ge f_\beta\right\}, \tag{2.56}$$

for some threshold f_β—equation (2.55) implies

$$\mathcal{X}_\beta(\mathbf{d}) = \left\{\mathbf{x}\colon \tau(\mathbf{x},\mathbf{d}) \le \tau_\beta\right\}, \tag{2.57}$$

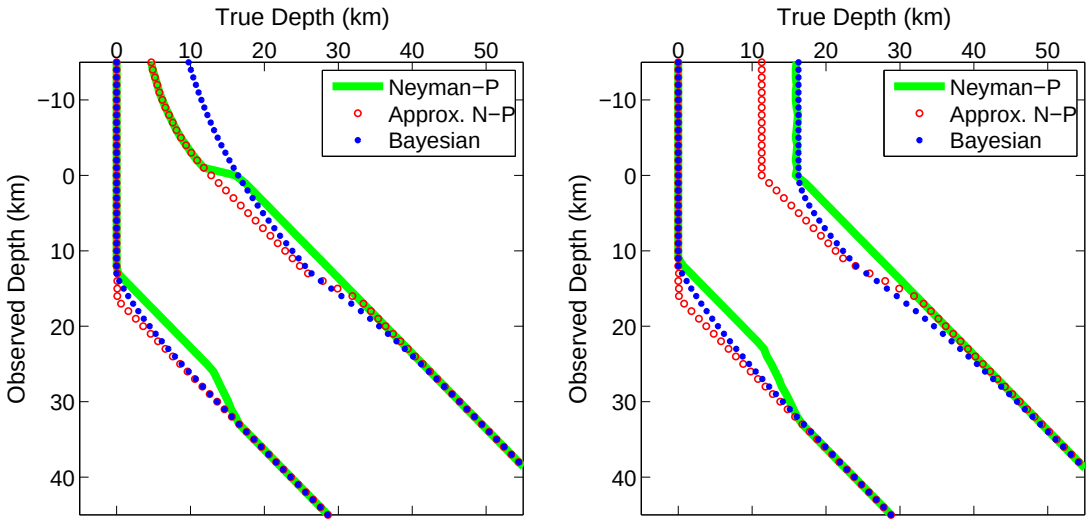

Figure 2.1: Confidence intervals on event depth in a fictitious problem involving a direct observation of depth, which is either Gaussian (left panel) or Laplace (right) distributed, with a variance of $(10 \text{ km})^2$ in each case. Three types of confidence limits (at $\beta = 90\%$) are shown, each as a function of the depth observation itself (vertical axis): exact Neyman-Pearson (green line), approximate Neyman-Pearson (red dots), and Bayesian (blue dots). For a given value of observed depth, the confidence interval of each type consists of the range of true depth between the confidence limits.

where the cut-off τ_β satisfies

$$\int_{\tau(\mathbf{x},\mathbf{d})\leq\tau_\beta} d\mathbf{x}\; e^{-\tau(\mathbf{x},\mathbf{d})} = \beta \int_{\tau(\mathbf{x},\mathbf{d})<\infty} d\mathbf{x}\; e^{-\tau(\mathbf{x},\mathbf{d})}. \tag{2.58}$$

This definition of τ_β offers an alternative approximation to $\widetilde{\tau}_\beta(\mathbf{x})$ of the exact Neyman-Pearson approach.

A numerical algorithm for computing quasi-Bayesian confidence regions shares the likelihood mapping step with the approximate Neyman-Pearson method, but replaces the Monte-Carlo simulation step with a much simpler and more efficient process we call *likelihood integration*. This process creates a histogram of the likelihood function values found in the mapping step. The histogram allows one to find approximate solutions of equation (2.58) for τ_β as a function of β. While the likelihood integration process is much faster computationally than Monte Carlo simulation, it does place some additional burden on the mapping step. The mapping must be sufficiently thorough to yield accurate approximations to the integrals in equation (2.58).

We make two more observations about approximate Neyman-Pearson and quasi-Bayesian confidence regions. First, neither method guarantees the confidence property, or even its weaker version in equation (2.37). Second, in the Gaussian/linear problem, the methods coincide with each other and with the exact Neyman-Pearson result.

2.4.3 Comparison in a simple situation

The difference between exact Neyman-Pearson, approximate Neyman-Pearson and quasi-Bayesian confidence regions in the presence of nonlinearity is illustrated in Figure 2.1. This is a "toy" problem involving one unknown parameter—event depth—and a single datum that is taken to

be a direct, unbiased estimate of the depth. We can state this inverse problem as

$$d = z + e, \tag{2.59}$$

where d represents the observed depth. We introduce nonlinearity into the problem in the form of the hard constraint that the true event depth (horizontal axes in the figure) cannot be negative, even though the observed depth (vertical axes) might be, such as when arrival times yield an "airquake" as the best-fitting solution for an event hypocenter. For any probability distribution of e that is symmetric about zero ($f[e] = f[-e]$), the maximum-likelihood estimate of z, subject to the positivity constraint, is

$$\hat{z} = \max(d, 0). \tag{2.60}$$

Intuitively, the positivity constraint should affect a confidence interval on depth when the observed depth is negative or near $z = 0$, and we can see from the figure that this is the case. The results in the left panel assumes that the error in the observed depth (e) has a Gaussian distribution ($p = 2$), while the right panel assumes that e has a Laplace distribution ($p = 1$).

The green lines in each panel of Figure 2.1 trace the confidence limits for exact Neyman-Pearson (N-P) confidence intervals, which possess the confidence property. The red dots are the approximate Neyman-Pearson confidence limits that result when the critical value τ_β is considered only for the maximum-likelihood estimate of depth (as per equation (2.51)). The blue dots mark the quasi-Bayesian confidence limits, which are actually proper Bayesian limits in this toy problem since no nuisance parameters are involved. Generally, the approximate Neyman-Pearson intervals are smaller than the Bayesian intervals for shallow and negative depth observations. Which one better represents the exact N-P intervals (green lines) depends on the type of error distribution used.

2.5 Examples for a Nevada Test Site Event

Figure 2.2 shows an example of our numerical confidence region algorithm applied to a Nevada Test Site (NTS) explosion at the Pahute Mesa testing area. The calculations were done using six Pn arrivals for the event, assuming Gaussian pick errors with know variance. The left panel shows the event-station geometry. The center panel shows the log-likelihood function mapped on an epicenter grid. The right panel shows the confidence regions, at three confidence levels, that result after performing the likelihood integration step to find τ_β. The confidence regions were computed with the assumption that the event focal depth is known (fixed). We see in the right panel that the epicentral confidence region for each confidence level is very close to elliptical in shape, indicating that nonlinearity of the travel-time forward model does not have a significant effect.

Figure 2.3 shows numerical confidence regions for the same Pahute Mesa event computed with various non-Gaussian pick error distributions. The confidence regions become less elliptical as the error distribution becomes increasingly less Gaussian (left to right).

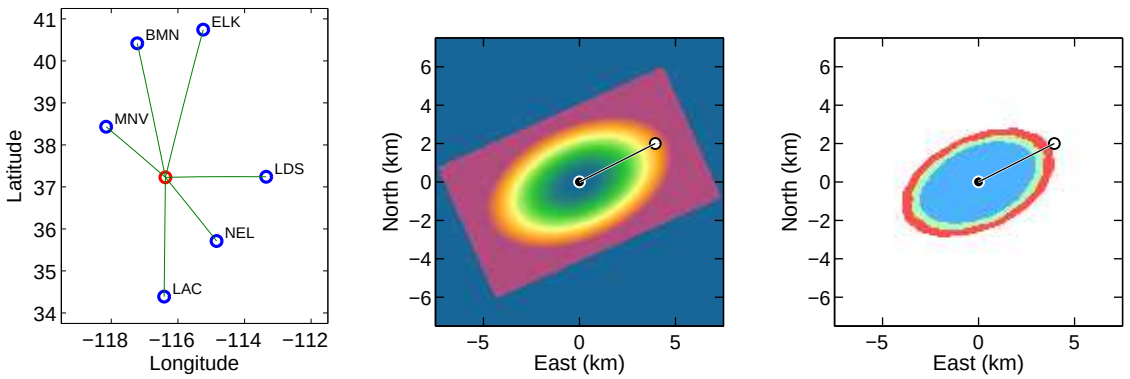

Figure 2.2: Log-likelihood function and confidence regions for a Pahute Mesa explosion, derived from 6 Pn arrivals with the pick error distribution assumed to be Gaussian ($p = 2$). *Left:* Schematic map showing the event epicenter (red dot) and station locations (blue dots). *Center:* Log-likelihood function mapped on an epicenter grid. *Right:* Epicenter confidence regions (determined from the likelihood function) for 90, 95 and 98% confidence (blue, green and red, respectively). In the right two panels, the black circle marks the maximum-likelihood estimate for the event epicenter, and the white circle is its GT0 epicenter (from Walter *et al.*, 2003).

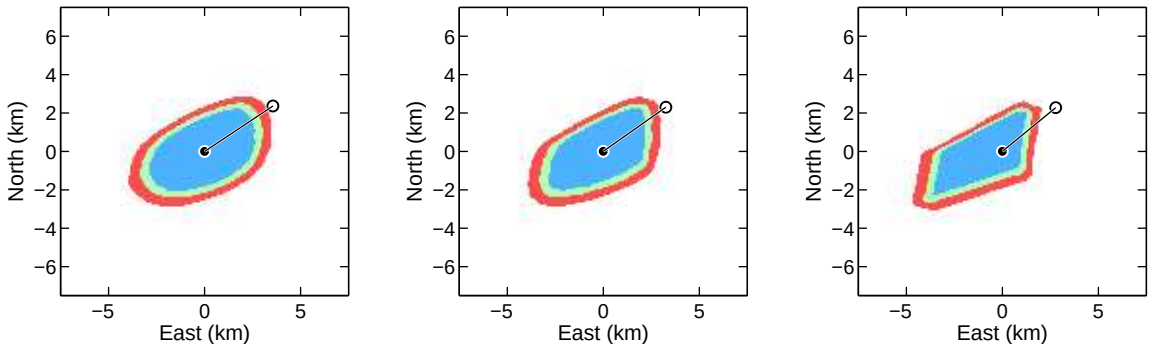

Figure 2.3: Numerical confidence regions for the same Pahute Mesa explosion as in Figure 2.2, computed with non-Gaussian error distributions: $p = 1.5$ (left), $p = 1.25$ (center) and $p = 1$ (right).

Chapter 3

Model Errors and the Joint Location/Calibration Inverse Problem

3.1 Model Errors

We extend equation (2.1) to include an additional term, c_i, which represents a correction to the travel-time predicted by the function T_i:

$$d_i = T_i(\mathbf{x}) + t + c_i + e_i, \quad i = 1, \ldots, n. \tag{3.1}$$

Assuming they are not known, the travel-time corrections c_i play the role of prediction errors, or *model errors*, in the event location problem. Defining a correction vector,

$$\mathbf{c} = \begin{pmatrix} c_1 & c_2 & \cdots & c_n \end{pmatrix}^T, \tag{3.2}$$

we can write (3.1) in vector notation as

$$\mathbf{d} = \mathbf{T}(\mathbf{x}) + \mathbf{1}t + \mathbf{c} + \mathbf{e}. \tag{3.3}$$

Equation (3.1) is a hopelessly ill-posed inverse problem unless we impose some sort of prior constraints on the c_i. A common way of doing this is to assume the model and pick errors are both Gaussian random variables and combine them as a single error. Each data variance, σ_i^2, is then reinterpreted as the variance of the error sum, $c_i + e_i$. This approach ignores the different natures of pick and model errors and thus must be considered ad hoc. For example, there is no physical basis for assuming that the model errors for different station/phase pairs are necessarily statistically independent, as is usually assumed for pick errors. If we consider the model errors as distinct from the pick errors, we have the opportunity to constrain them with a less restrictive probability model.

A more general approach is to retain the c_i as explicit unknowns in the inverse problem and constrain them with prior information in the form of a prior probability density function. We allow this p.d.f. to depend on the target-event hypocenter and denote it as $f[\mathbf{c}; \mathbf{x}]$. In the maximum-likelihood formulation, the prior p.d.f. on $\mathbf{c}$ is included by augmenting the penalty function for the arrival data with a prior penalty function for $\mathbf{c}$, given by

$$\Phi_c(\mathbf{c}; \mathbf{x}) = -\log f[\mathbf{c}; \mathbf{x}]. \tag{3.4}$$

Thus, the likelihood function for the single-event location problem, including model errors, becomes

$$\Lambda(\mathbf{x}, t, \mathbf{c}, \sigma; \mathbf{d}) = n \log K(p) + \sum_{i=1}^{n} \log \nu_i + n \log \sigma + \frac{1}{p\sigma^p} \Psi(\mathbf{x}, t, \mathbf{c}; \mathbf{d}) + \Phi_c(\mathbf{c}; \mathbf{x}), \tag{3.5}$$

where the data-misfit function is now

$$\Psi(\mathbf{x}, t, \mathbf{c}; \mathbf{d}) = \sum_{i=1}^{n} |d_i - T_i(\mathbf{x}) - t - c_i|^p / \nu_i^p. \tag{3.6}$$

Given this, the theory of Section 2.3 applies to the new problem almost as written, but with $\mathbf{c}$ adding to the list of nuisance parameters. The maximum-likelihood solution now satisfies

$$\Lambda(\widehat{\mathbf{x}}, \widehat{t}, \widehat{\mathbf{c}}, \widehat{\sigma}; \mathbf{d}) = \min_{\mathbf{x}, t, \mathbf{c}, \sigma} \Lambda(\mathbf{x}, t, \mathbf{c}, \sigma; \mathbf{d}). \tag{3.7}$$

The reduced penalty function is now given by

$$\widetilde{\Lambda}(\mathbf{x}; \mathbf{d}) = \min_{t, \mathbf{c}, \sigma} \Lambda(\mathbf{x}, t, \mathbf{c}, \sigma; \mathbf{d}). \tag{3.8}$$

The test statistic for hypocenter confidence regions is still expressed by equation (2.33), and the numerical methods described in Section 2.4 still apply. In the case of Monte Carlo simulation, synthetic data are computed as

$$\mathbf{d}^{\mathsf{mc}} = \mathbf{T}(\widehat{\mathbf{x}}) + \mathbf{1}\widehat{t} + \widehat{\mathbf{c}} + \mathbf{e}^{\mathsf{mc}}(\widehat{\sigma}). \tag{3.9}$$

The difficult issue, which was the focus of this project, is how to choose the prior penalty function $\Phi_c(\mathbf{c}; \mathbf{x})$. The approach we took was to consider Φ_c to be the *posterior* negative log-likelihood that derives from a calibration analysis.

3.2 Calibration and the Prior Likelihood on Corrections

We assume that calibration is performed with seismic arrival-time data observed from a set of m calibration events. In analogy with equation (3.1), then, the calibration problem can be expressed as

$$d_{ij} = T_{ij}(\mathbf{x}_j) + t_j + c_{ij} + e_{ij}, \quad i = 1, \ldots, n_j, \ j = 1, \ldots, m. \tag{3.10}$$

We do not assume that all observable station/phase combinations have actually been observed for every calibration event, or for the target event. Therefore, the particular station/phase indexed by a given value of i may differ between events. In vector notation, equation (3.10) becomes

$$\mathbf{d}_j = \mathbf{T}_j(\mathbf{x}_j) + \mathbf{1}t_j + \mathbf{c}_j + \mathbf{e}_j, \quad j = 1, \ldots, m, \tag{3.11}$$

where, for a given j, the vectors have n_j components.

We also do not assume that the calibration events are necessarily GT0 events having perfectly known hypocenters and origin times. In general, therefore, the unknowns in the calibration inverse problem are the m event locations, $(\mathbf{x}_j, t_j)$, and $\sum_{j=1}^{m} n_j$ path travel-tine corrections, c_{ij}. With both correction and location parameters both unknown, the calibration problem may also be referred to as the *joint location/calibration* problem, although we reserve that term for

the situation in which a target event location is involved. Additionally, one may refer to the problem as *multiple-event location*.

As in the single-event location problem for a target event, we assume that the observational errors in the calibration problem, e_{ij}, have a generalized Gaussian distribution of order p. Unlike the target-event problem, however, we take the standard errors, σ_{ij}, to be known completely. The data misfit function for the jth calibration events is thus given by

$$\Psi_j(\mathbf{x}_j, t_j, \mathbf{c}_j; \mathbf{d}_j) = \sum_{i=1}^{n_j} |d_{ij} - T_{ij}(\mathbf{x}) - t_j - c_{ij}|^p / \sigma_{ij}^p. \tag{3.12}$$

3.2.1 Parameterization of corrections

To link the target-event location problem of equation (3.1) and calibration problem of equation (3.10), we must represent the path travel-time corrections, c_i and c_{ij}, with a common parameterization. Let us consider four such parameterizations in their order of increasing complexity. To facilitate this discussion, we will refer to the target event with the index $j = 0$, allowing equation (3.1) to be expressed by equation (3.10). Therefore, the corrections c_i are alternatively denoted as c_{i0}, and the target-event location parameters $(\mathbf{x}, t)$ as $(\mathbf{x}_0, t_0)$.

First, in the "basic" multiple-event location problem (e.g. Jordan and Sverdrup, 1981; Pavlis and Booker, 1983), travel-time corrections are assumed to be common to the events but different between station/phase combinations. If there are ℓ unique combinations, the relevant correction parameters are time terms, a_k, $k = 1, \ldots, \ell$, such that

$$c_{ij} = a_{k_{ij}}, \tag{3.13}$$

where k_{ij} denotes the station/phase pair associated with the ith observation for the jth event.

A second possible parameterization, introduced by Schultz *et al.* (1998), comprises a correction "surface," or function, specific to each station/phase. Each surface, $a_k(\mathbf{x})$, defines a continuous dependence of the travel-time correction for an event on the event hypocenter. In this case we have

$$c_{ij} = a_{k_{ij}}(\mathbf{x}_j). \tag{3.14}$$

A third example of a correction parameterization is the vector "mislocation function," $\mathbf{a}(\mathbf{x})$, introduced by Rodi *et al.* (2005) and Murphy *et al.* (2005). In this case,

$$c_{ij} = \mathbf{p}_{ij}(\mathbf{x}_j, \mathbf{y}_{ij}) \cdot \mathbf{a}(\mathbf{x}_j) + \mathbf{q}_{ij}(\mathbf{x}_j, \mathbf{y}_{ij}) \cdot \mathbf{a}(\mathbf{y}_{ij}), \tag{3.15}$$

where $\mathbf{y}_{ij}$ is the station position associated with the ijth observation, and where $\mathbf{p}_{ij}$ and $\mathbf{q}_{ij}$ are slowness vectors (travel-time gradients) at the event and station locations, respectively. Unlike equations (3.13) and (3.14), this parameterization ensures source-receiver reciprocity (c_{ij} is the same if $\mathbf{x}_j$ and $\mathbf{y}_{ij}$ are swapped).

Finally, we can relate travel-time corrections to 3-D velocity anomalies in the Earth. Under the linear approximation to the velocity dependence of travel times, the *tomographic* parameterization of travel-time corrections is expressed as

$$c_{ij} = \int d\mathbf{x}\; b_{ij}(\mathbf{x}; \mathbf{x}_j)\, \delta u(\mathbf{x}), \tag{3.16}$$

where $b_{ij}(\mathbf{x}; \mathbf{x}_j)$ is a travel-time sensitivity kernel and $\delta u(\mathbf{x})$ is the slowness difference between the real Earth and the reference Earth model used for evaluating the travel-time function T_{ij}.

In this case, the slowness perturbation $\delta u(\mathbf{x})$ serves as the common parameterization for travel-time corrections.

Presuming that continuous parameter functions—such as $a_k(\mathbf{x})$, $\mathbf{a}(\mathbf{x})$ and $\delta u(\mathbf{x})$ above—are discretized on a spatial grid, we can generalize these specific parameterizations and others as

$$c_{ij} = \mathbf{b}_{ij}(\mathbf{x}_j)^T \mathbf{u}, \quad i = 1, \ldots, n_j, \ j = 0, 1, \ldots, m, \tag{3.17}$$

where the vector $\mathbf{u}$ contains a discrete set of parameters for generating travel-time corrections. It will be convenient to write equation (3.17) alternatively as

$$\mathbf{c}_j = \mathbf{B}_j(\mathbf{x}_j)\mathbf{u}, \quad j = 0, 1, \ldots, m, \tag{3.18}$$

where each *sensitivity* matrix $\mathbf{B}_j(\mathbf{x})$ is given by

$$\mathbf{B}_j(\mathbf{x}_j) = \begin{pmatrix} \mathbf{b}_{1j}(\mathbf{x}_j)^T \\ \mathbf{b}_{2j}(\mathbf{x}_j)^T \\ \vdots \\ \mathbf{b}_{n_j j}(\mathbf{x}_j)^T \end{pmatrix}. \tag{3.19}$$

For $j = 0$ (target event) we also write (3.18) as

$$\mathbf{c} = \mathbf{B}(\mathbf{x})\mathbf{u}. \tag{3.20}$$

The unknowns in the calibration inverse problem can now be identified as the parameter vector $\mathbf{u}$ and calibration event locations $(\mathbf{x}_j, t_j)$, $j = 1, \ldots, m$.

3.2.2 Penalty function for the calibration problem

A negative log-likelihood function for the calibration problem is given by

$$\begin{aligned} \Lambda_{\mathsf{cal}}(\mathbf{u}, \mathbf{x}_1, t_1, \ldots, \mathbf{x}_m, t_m; \mathbf{d}_1, \ldots, \mathbf{d}_m) = \text{const} \\ + p^{-1} \sum_{j=1}^{m} \Psi_j(\mathbf{x}_j, t_j, \mathbf{B}_j(\mathbf{x}_j)\mathbf{u}; \mathbf{d}_j) \\ + \sum_{j=1}^{m} \Phi_j(\mathbf{x}_j, t_j) + \Phi_u(\mathbf{u}). \end{aligned} \tag{3.21}$$

The constant term ("const") is not elaborated because it does not depend on any of the problem unknowns. The first summation includes the data-misfit functions for the calibration arrival data. These are given in equation (3.12), with equation (3.18) substituting for $\mathbf{c}_j$. Each term of the second summation represents a prior penalty function for the location parameters of a calibration event. The final term is a prior penalty function for the correction parameters. An appropriate choice for Φ_u would depend on the physical nature of the parameters in $\mathbf{u}$.

For this project we assumed that prior probability distributions on $\mathbf{x}$ and t are of the generalized Gaussian type. Further, the epicentral parameters (latitude θ and longitude ϕ), focal depth z and origin time were assumed to be independent. Therefore, the prior penalty function for a calibration event is given by (omitting the subscript j)

$$\Phi(\mathbf{x}, t) = \mathsf{const} + \frac{1}{p_\Delta} \left| \frac{\Delta(\theta, \phi; \theta_{\mathsf{pri}}, \phi_{\mathsf{pri}})}{\sigma_\Delta} \right|^{p_\Delta} + \frac{1}{p_z} \left| \frac{z - z_{\mathsf{pri}}}{\sigma_z} \right|^{p_z} + \frac{1}{p_t} \left| \frac{t - t_{\mathsf{pri}}}{\sigma_t} \right|^{p_t}, \tag{3.22}$$

where $(\theta_{\mathsf{pri}}, \phi_{\mathsf{pri}}, z_{\mathsf{pri}})$ is a prior hypocenter and t_{pri} a prior origin time for the event. The function Δ computes the spherical distance between two geographic points. Thus, σ_Δ is the standard error on the spherical distance of the event epicenter (θ, ϕ) from $(\theta_{\mathsf{pri}}, \phi_{\mathsf{pri}})$, while σ_z and σ_t are standard errors on depth and time. We note that the prior penalty function effects hard bounds on the parameters as the generalized Gaussian orders (p_Δ, p_z, p_t) tend toward infinity.

The maximum-likelihood solution of the calibration inverse problem is defined by minimizing the penalty function Λ_{cal} jointly with respect to the correction parameters, $\mathbf{u}$, and the location parameters $(\mathbf{x}_j, t_j)$ of the (non-GT0) calibration events. How this is best done algorithmically depends on the type of parameters in $\mathbf{u}$. In designing a minimization algorithm, it is useful to recognize, by examining the expression for Λ_{cal} in (3.21), that, for any fixed value of $\mathbf{u}$, minimization with respect to the event parameters decouples across events and equates to performing single-event location independently on each of the calibration events.

3.2.3 The posterior on c

A prior penalty function on the travel-time corrections $\mathbf{c}$, needed for locating the target event, is provided by a posterior penalty function on $\mathbf{c}$ that results from solving the calibration problem. In the maximum-likelihood framework, we can define the latter in a quasi-Bayesian sense, whereby unknown parameters other than $\mathbf{c}$ are projected from the problem by minimization. However, since $\mathbf{c}$ itself is not an explicit parameter in the calibration problem, but rather is a mapping of $\mathbf{u}$, projection of other parameters is accomplished by *constrained* minimization. Thus, we can state formally

$$\Phi_c(\mathbf{c};\mathbf{x}) = \min_{\mathbf{u}:\mathbf{B}(\mathbf{x})\mathbf{u}=\mathbf{c}} \min_{\mathbf{x}_1,t_1,\ldots,\mathbf{x}_m,t_m} \Lambda_{\mathsf{cal}}(\mathbf{u},\mathbf{x}_1,t_1,\ldots,\mathbf{x}_m,t_m;\mathbf{d}_1,\ldots,\mathbf{d}_m). \tag{3.23}$$

The minimization on the right-hand-side yields a solution of the calibration problem with $\mathbf{u}$ constrained as $\mathbf{B}(\mathbf{x})\mathbf{u} = \mathbf{c}$. Substituting from (3.21), we also have

$$\Phi_c(\mathbf{c};\mathbf{x}) = \mathsf{const} + \min_{\mathbf{u}:\mathbf{B}(\mathbf{x})\mathbf{u}=\mathbf{c}} \Big\{ \sum_{j=1}^{m} \min_{\mathbf{x}_j,t_j} \Big[p^{-1}\Psi_j(\mathbf{x}_j,t_j,\mathbf{B}_j(\mathbf{x}_j)\mathbf{u};\mathbf{d}_j) + \Phi_j(\mathbf{x}_j,t_j) \Big] + \Phi_u(\mathbf{u}) \Big\}. \tag{3.24}$$

One way to interpret this equation is that the minimization over the calibration-event location parameters transforms the prior log-likelihood on $\mathbf{u}$, $-\Phi_u(\mathbf{u})$, into a posterior log-likelihood on $\mathbf{u}$. The constrained minimization over $\mathbf{u}$ then transforms the posterior om $\mathbf{u}$ into a posterior on $\mathbf{c}$.

Computing the function Φ_c for use in locating target events is problematical. In general, Φ_c does not admit a closed form expression. The alternative of representing Φ_c in tabular form is not appealing since a grid of values in n-dimensional $\mathbf{c}$-space would be unmanageably large. Additional difficulty arises from the dependence of Φ_c on $\mathbf{x}$. The $\mathbf{x}$-dependence also defies a closed form representation and would add to the dimensionality of a tabular representation.

3.3 Joint Location/Calibration Uncertainty Analysis

A confidence region on the target-event hypocenter $\mathbf{x}$ is obtained by mapping the function $\widetilde{\Lambda}(\mathbf{x})$ in hypocenter space, which in turn requires the evaluation of $\Phi_c(\mathbf{c};\mathbf{x})$ for each grid point $\mathbf{x}$ (see equations (3.5) and (3.8)). If one could evaluate equation (3.23) in closed form, it would be

possible to separate the computations involved in calibration and target event location. In general, however, this is not possible without invoking some approximation. In an attempt to avoid approximations, this project has investigated an approach that bypasses the computation of Φ_c by combining the target-event location and calibration problems into a joint inverse problem. This means substituting equation (3.23) into (3.8) and setting $\mathbf{c} = \mathbf{Bu}$, obtaining (after some rearranging)

$$\widetilde{\Lambda}(\mathbf{x};\mathbf{d}) = \min_{\mathbf{u}} \Big\{ \min_{t,\sigma} \Lambda(\mathbf{x}, t, \mathbf{B}(\mathbf{x})\mathbf{u}, \sigma; \mathbf{d}) + \min_{\mathbf{x}_1,t_1,\ldots,\mathbf{x}_m,t_m} \Lambda_{\mathsf{cal}}(\mathbf{u}, \mathbf{x}_1, t_1, \ldots, \mathbf{x}_m, t_m; \mathbf{d}_1, \ldots, \mathbf{d}_m) \Big\} \tag{3.25}$$

or

$$\widetilde{\Lambda}(\mathbf{x};\mathbf{d}) = \mathsf{const} + \min_{\mathbf{u}} \Big\{ \min_{\sigma} \Big[n \log \sigma + p^{-1} \sigma^{-p} \min_{t} \Psi(\mathbf{x}, t, \mathbf{B}(\mathbf{x})\mathbf{u}; \mathbf{d}) \Big] + \sum_{j=1}^{m} \min_{\mathbf{x}_j,t_j} \Big[p^{-1} \Psi_j(\mathbf{x}_j, t_j, \mathbf{B}_h(\mathbf{x}_j)\mathbf{u}; \mathbf{d}_j) + \Phi_j(\mathbf{x}_j, t_j) \Big] \Big\}. \tag{3.26}$$

The computational burden resulting from merging the location and calibration problems is apparent. To evaluate $\widetilde{\Lambda}(\mathbf{x};\mathbf{d})$ for one value of $\mathbf{x}$ (i.e. one point in a likelihood map) requires the minimization of a penalty function with respect to all parameters except $\mathbf{x}$, including the calibration parameters ($\mathbf{u}$) and calibration event locations ($\mathbf{x}_1$, t_1, ..., $\mathbf{x}_m$, t_m). That is, a large inverse problem must be solved repeatedly in order to calculate a confidence region on the target event hypocenter. The practicality of the joint location/calibration approach will depend on the nature of the parameterization of travel-time corrections.

3.4 Examples for Basic Multiple-Event Location

Computing confidence regions on a target event as part of a joint location/calibration analysis is a feasible task for the basic multiple-event location problem. In this problem, the travel-time corrections comprise a simple time term for each station/phase combination, as discussed in Section 3.2.1 (see equation (3.13)). We have implemented joint inversion uncertainty analysis for basic multiple-event location as part of a location program, GMEL (Grid-search Multiple-Event Location), developed under this and previous projects. The GMEL algorithm is described by Rodi *et al.* (2002) and Rodi (2006).

Figure 3.1 shows confidence regions for the same Pahute Mesa event used earlier (Figures 2.2 and 2.3), but now including the effects of model errors, which are equated to the uncertainty in the station/phase time terms estimated from calibration events. We used 32 other explosions at Pahute Mesa and Rainier Mesa as the calibration events. Only one of the calibration events, a relatively well-recorded event at Rainier Mesa (16 Pn arrivals), was assigned a finite ground-truth level. The three panels in the figure show the resulting confidence regions under three assumptions about the GT level of that calibration event: GT0, GT2 or GT5 (at 90% confidence). We see that the confidence regions, which now take model errors into account, are larger than when model errors are assumed to be zero (Figure 2.2), and grow as the uncertainty in the GT calibration event location is increased. Figure 3.2 repeats the GT0 case with various non-Gaussian pick error distribution: $p = 1.5$, 1.25 and 1 (left, center, right, respectively).

The confidence regions shown in Figures 3.1 and 3.2 took between 5 and 30 CPU minutes each to compute on a high-end Linux workstation, the ones for smaller p taking the longest. In

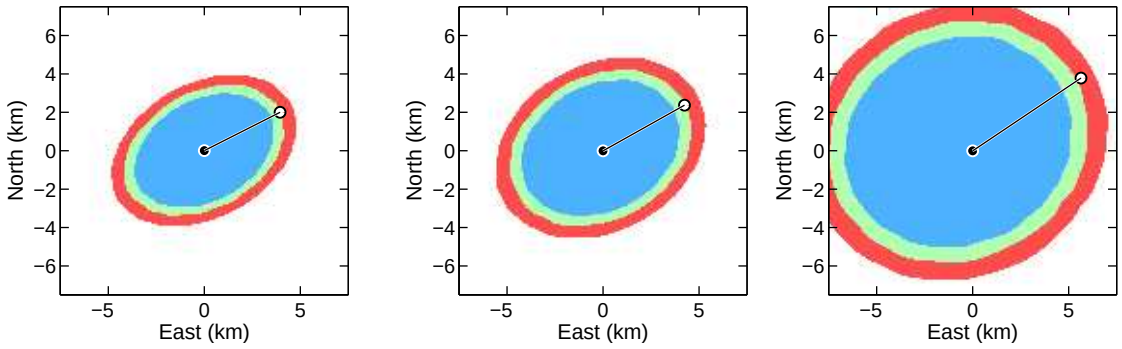

Figure 3.1: Confidence regions for the Pahute Mesa target event shown in Figure 2.2, but accounting for uncertainty in travel-time corrections (model errors). The corrections were constrained by 32 calibration events (other NTS explosions) with one of them assigned a finite GT level: GT0 (left), GT2 (center), or GT5 (right). The pick error distribution was assumed to be Gaussian ($p = 2$).

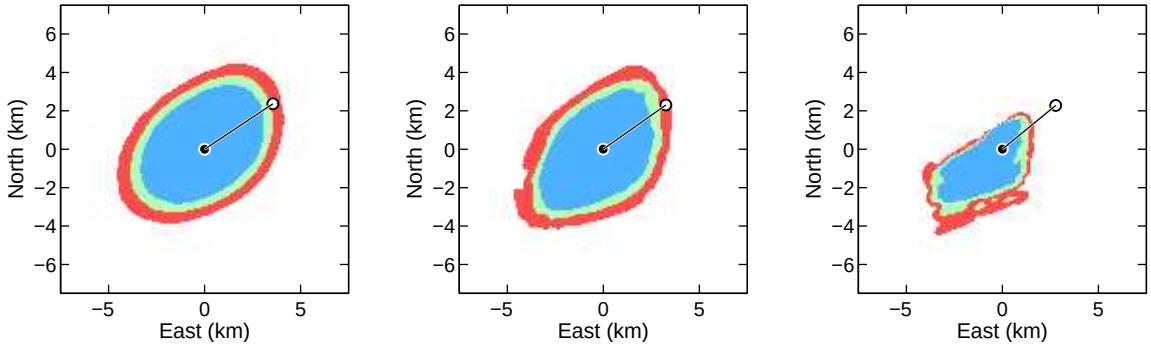

Figure 3.2: Confidence regions for the same Pahute Mesa event, accounting for model errors as constrained by 32 calibration events (see Figure 3.1). The GT calibration event is assumed to be GT0 and the pick error distribution is non-Gaussian: $p = 1.5$ (left), $p = 1.25$ (center), and $p = 1$ (right).

at least one case (right panel in Figure 3.2) the likelihood function was not mapped sufficiently well to compute accurate values of τ_β, meaning even more computation was needed. To apply the joint inversion approach with tomographic corrections, therefore, would be prohibitive since each point on a likelihood grid would require performing a full 3D tomography in conjunction with calibration event relocation.

Chapter 4

Two-Stage Uncertainty Analysis

To be computationally practical, in more than the simplest cases, the calculation of calibration uncertainty must be performed separately from target-event location. This, in fact, is the traditional practice; the challenge is to accomplish it rigorously and efficiently. In the previous chapter we mentioned that one way to accomplish the separation was by tabulating a prior likelihood for the correction vector $\mathbf{c}$ as the output of calibration analysis, and then using the table in the subsequent location of a target event. The difficulty here is that the table would be too large; e.g. 10^{14} numbers for 20 corrections sampled at five values each. This chapter presents a more efficient separation scheme, based on a Gaussian representation of correction uncertainty.

4.1 Gaussian (Quadratic) Approximation

Our two-stage scheme approximates the prior penalty function $\Phi_c(\mathbf{c};\mathbf{x})$ as quadratic in $\mathbf{c}$. Allowing the quadratic function to be different for different $\mathbf{x}$, we write

$$\Phi_c(\mathbf{c};\mathbf{x}) \approx \Phi_c(\mathbf{c}^*(\mathbf{x});\mathbf{x}) + \frac{1}{2}(\mathbf{c}-\mathbf{c}^*(\mathbf{x}))^T\mathbf{V}_c(\mathbf{x})^{-1}(\mathbf{c}-\mathbf{c}^*(\mathbf{x})), \tag{4.1}$$

where $\mathbf{V}_c(\mathbf{x})$ is a symmetric matrix and $\mathbf{c}^*(\mathbf{x})$ is a stationary point of Φ_c:

$$\nabla_{\mathbf{c}}\Phi_c(\mathbf{c}=\mathbf{c}^*(\mathbf{x});\mathbf{x}) = 0. \tag{4.2}$$

When $\mathbf{V}_c(\mathbf{x})$ is positive definite—implying Φ_c is concave near $\mathbf{c}=\mathbf{c}^*(\mathbf{x})$—the quadratic approximation corresponds to the assumption that $\mathbf{c}$ has a Gaussian prior probability distribution with mean $\mathbf{c}^*(\mathbf{x})$ and covariance matrix $\mathbf{V}_c(\mathbf{x})$.

We now outline algorithms we developed for calculating $\mathbf{c}^*$ and $\mathbf{V}_c$ in a calibration analysis (Stage 1) and using them in locating a target event (Stage 2).

4.2 Stage 1: Calibration

Although the vector $\mathbf{c}^*$ and matrix $\mathbf{V}_c$ depend on $\mathbf{x}$, we will not show this dependence in this section for the sake of notational clarity.

Let $\widehat{\mathbf{u}}$ be the maximum-likelihood solution for the correction parameter vector $\mathbf{u}$; i.e., $\widehat{\mathbf{u}}$ minimizes Λ_{cal} in equation (3.21) in conjunction with maximum-likelihood estimates of the calibration event locations. We set $\mathbf{c}^*$ to

$$\mathbf{c}^* = \mathbf{B}\widehat{\mathbf{u}}, \tag{4.3}$$

where $\mathbf{B}$ is the sensitivity matrix for the target-event corrections. Since Λ_{cal} is a minimum at $\mathbf{u} = \widehat{\mathbf{u}}$, equations (3.23) and (4.3) imply that Φ_c is a minimum at $\mathbf{c} = \mathbf{c}^*$. In the absence of bounds on $\mathbf{c}$, we infer that $\mathbf{c} = \mathbf{c}^*$ is a stationary point of Φ_c.

Our algorithm for calculating $\mathbf{V}_c$ is a perturbation technique. The algorithm is adapted from a method developed by Rodi and Myers (2007) for computing travel-time covariances (their "implicit" method); we refer the reader to that paper for details about the technique. Applied here, the algorithm solves a perturbed calibration problem for each of the n corrections, c_i, associated with the target event. Each problem requires minimizing an augmented penalty function given by

$$\Lambda_{\mathsf{cal}}^{(i)}(\mathbf{u}, \dots) = \Lambda_{\mathsf{cal}}(\mathbf{u}, \dots) + \frac{1}{2\rho}(\mathbf{c}^* + \epsilon\, \mathbf{n}^{(i)} - \mathbf{B}\mathbf{u})^T(\mathbf{c}^* + \epsilon\, \mathbf{n}^{(i)} - \mathbf{B}\mathbf{u}). \tag{4.4}$$

The vector $\mathbf{n}^{(i)}$ is the ith column of the identity matrix, containing unity in its ith element and zero elsewhere. The values of the scalars ϵ and ρ control the numerical stability of the algorithm. Let $\widehat{\mathbf{u}}^{(i)}$ denote the solution of the perturbed calibration problem and define the residual vector

$$\mathbf{r}^{(i)} = \mathbf{c}^* + \epsilon \mathbf{n}^{(i)} - \mathbf{B}\widehat{\mathbf{u}}^{(i)}. \tag{4.5}$$

Calculating this vector for each of the n perturbed problems we can construct a residual *matrix* as

$$\mathbf{R} = \left(\mathbf{r}^{(1)}\ \mathbf{r}^{(2)}\ \cdots\ \mathbf{r}^{(n)}\right). \tag{4.6}$$

Then $\mathbf{V}_c$ is obtained as

$$\mathbf{V}_c = \rho\,(\epsilon\,\mathbf{R}^{-1} - \mathbf{I}), \tag{4.7}$$

where $\mathbf{I}$ is the $n \times n$ identity matrix.

We note that the perturbation method is exact if Λ_{cal}, minimized with respect to the $(\mathbf{x}_j, t_j)$, is a quadratic function of $\mathbf{u}$, in which case Φ_c is quadratic in $\mathbf{c}$. In general, however, the method can be viewed as a technique for approximating the Hessian matrix of Φ_c or fitting a quadratic function to Φ_c.

4.2.1 Dependence on x

The quantities $\mathbf{c}^*$ and $\mathbf{V}_c$ will depend on the target hypocenter $\mathbf{x}$ exactly when $\mathbf{B}$, used in equations (4.3) and (4.4), depends on $\mathbf{x}$. The above algorithms for $\mathbf{c}^*$ and $\mathbf{V}_c$ are still valid when this dependence occurs, but would have to be applied for each value of $\mathbf{x}$ of interest.

4.3 Stage 2: Location of a Target Event

Under the Gaussian approximation of Φ_c, the negative log-likelihood function for the target-event location problem, equation (3.5), becomes

$$\begin{aligned}\Lambda(\mathbf{x}, t, \mathbf{c}, \sigma; \mathbf{d}) = \mathsf{const} + n \log \sigma + \frac{1}{p\sigma^p}\Psi(\mathbf{x}, t, \mathbf{c}; \mathbf{d}) \\ + \frac{1}{2}(\mathbf{c} - \mathbf{c}^*(\mathbf{x}))^T \mathbf{V}_c(\mathbf{x})^{-1}(\mathbf{c} - \mathbf{c}^*(\mathbf{x})),\end{aligned} \tag{4.8}$$

where Ψ is given by equation (3.6). In general, $\mathbf{V}_c(\mathbf{x})$ is a full (non-diagonal) matrix. We have designed an algorithm to perform single-event location with a penalty function of this form. To date, the algorithm is only implemented for the case in which $\mathbf{c}^*$ and $\mathbf{V}_c$ do *not* depend on $\mathbf{x}$, and we describe this case first. This situation occurs with the time-term parameterization of travel-time correction discussed in Section 3.2.1 (equation (3.13)), i.e. the basic multiple-event location problem.

The algorithm minimizes Λ in a hierarchical fashion with the solution for $\mathbf{c}$ updated in an outer loop and the solution for the other parameters found inside the loop. Let us define the reduced penalty function

$$\bar{\Lambda}(\mathbf{c};\mathbf{d}) = \min_{\mathbf{x},t,\sigma} \Lambda(\mathbf{x},t,\mathbf{c},\sigma;\mathbf{d}). \tag{4.9}$$

The outer loop of the algorithm minimizes $\bar{\Lambda}$ with a nonlinear conjugate gradients (NLCG) technique (see, for example, Rodi and Mackie, 2001). NLCG is an iterative method that produces a sequence of estimates $\mathbf{c}_0$, $\mathbf{c}_1$, ... that converges to the maximum-likelihood solution, $\widehat{\mathbf{c}}$.

The kth step of the NLCG iteration requires the evaluation of $\bar{\Lambda}(\mathbf{c}_k)$ and the gradient of $\bar{\Lambda}$. These quantities are obtained by performing the minimization in equation (4.9) with $\mathbf{c}$ fixed to $\mathbf{c}_k$. That is, single-event location of the target event is performed with travel-time corrections fixed to their current estimate. Our implementation performs this task with the grid-search even location program GMEL (e.g. Rodi, 2006). The result of the grid-search minimization, $(\mathbf{x}_k, t_k, \sigma_k)$, converges to the maximum-likelihood solution as the NLCG iteration converges.

4.3.1 Dependence on x

The single-event location algorithm just described could account for the dependence of $\mathbf{c}^*$ and $\mathbf{V}_c$ on $\mathbf{x}$ in the grid-search location step, if these quantities were computed on a sufficiently fine grid in $\mathbf{x}$-space. A more practical alternative may be to embed the NLCG loop inside a higher level loop that updates the values of $\mathbf{c}^*$ and $\mathbf{V}_c$. The first NLCG solution would use values evaluated at an initial hypocenter $\mathbf{x}_0$. The quantities could then be re-evaluated at the resulting location solution, $\widehat{\mathbf{x}}$, for use in a second NLCG iteration. And so on. The efficiency and convergence properties of such an approach are a topic for future research.

4.4 Examples of Two-Stage Approach

We implemented the two-stage approach for the problem of basic multiple-event location. We show examples obtained with the same NTS calibration events, Pahute Mesa target event, and Pn arrival-time data used in previous examples. The event-station geometry was shown in Figure 2.2.

4.4.1 Calibration

Using the perturbation algorithm described in Section 4.2, we computed the 6×6 covariance matrix $\mathbf{V}_c$ on the travel-time corrections involved with the target event, based on data from 32 calibration events. One of these calibration events was assigned a finite GT level on its location, as described in Section 2.5.

Table 4.1 lists the correction standard deviations at two of the stations for various GT assignments. In addition to varying the epicenter error as before (0, 2 or 5 km), we also varied

Table 4.1: Correction Standard Deviation (sec) at Stations MNV and LAC

	Station MNV			Station LAC		
Or. time	*Epicenter error (km)*			*Epicenter error (km)*		
error	0	2	5	0	2	5
0	.13	.17	.28	.17	.20	.30
1	.44	.44	.49	.45	.47	.53
5	.58	.59	.62	.60	.61	.66

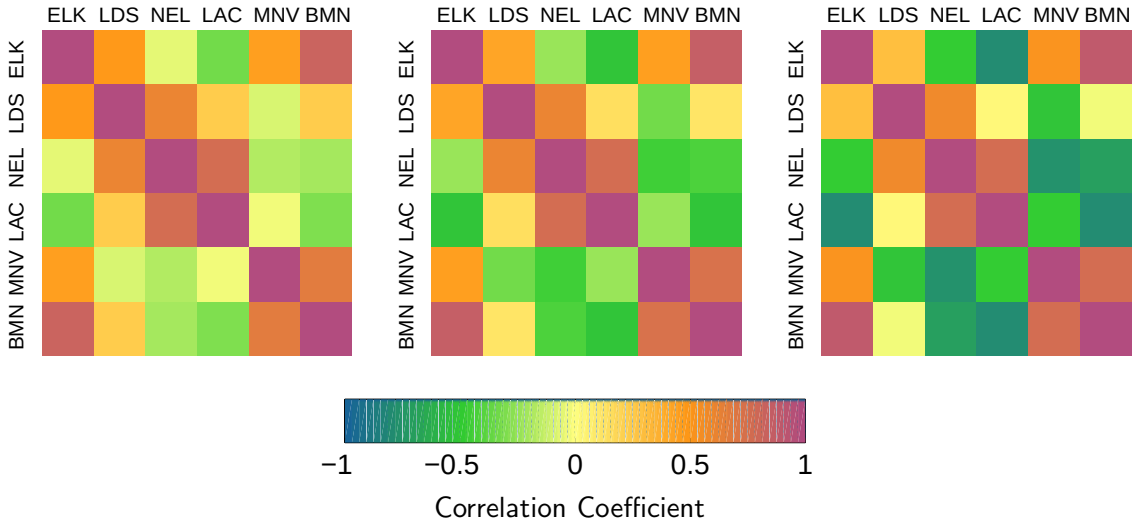

Figure 4.1: The correlation matrix for travel-time corrections at six stations recording the Pahute Mesa target event. The three panels correspond to different GT levels assigned to the epicenter of the ground-truth calibration event at Rainier Mesa. The origin time error of the GT event was set to zero.

the origin time GT level (0, 1 or 5 sec). The table shows that the uncertainty in the estimated travel-time corrections depends primarily on the origin-time error of the GT event.

Figure 4.1 displays the 6×6 correlation matrix of the estimated travel-time corrections for the three of nine GT cases in which the origin time error of the GT event is zero. Figure 4.2 shows the three cases in which the origin time error is 1 second. We see that when the origin-time error of the ground-truth event is not zero, the estimated corrections are strongly correlated. Not shown are the cases where the GT origin time error is 5 sec; then the correlations are even closer to 1.0 than in Figure 4.2.

4.4.2 Confidence regions on the target event epicenter

We generated numerical confidence regions for the Pahute Mesa target event using the travel-time correction means and covariances obtained from the calibration stage. The likelihood maps were calculated with the NLCG/grid-search location algorithm described in Section 4.3. We confirmed numerically that confidence regions on the epicenter do *not* depend on the origin-time GT level assumed in calibration, even though the covariance matrices did.

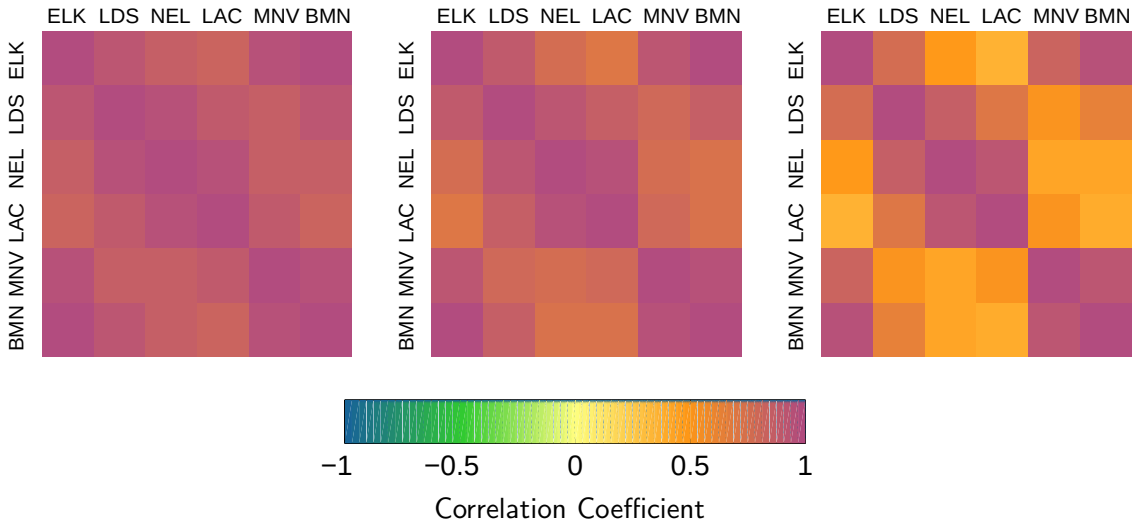

Figure 4.2: Same as Figure 4.1 except the origin time error of the GT calibration event is 1 second.

Figures 4.3 and 4.4 show confidence regions obtained with our new two-stage approach for the same cases shown for the joint approach in Figures 3.1 and 3.2. The differences from the joint location/calibration results are small except for the GT5 case. This may reflect the finer and more accurate likelihood mapping possible in the two-stage approach owing to its greater efficiency. The likelihood mapping with the two-stage approach takes only a few seconds of CPU time, rather than several minutes as in the joint approach. Another possible cause of the differences between Figures 3.2 and 4.4 is that the Gaussian approximation to the posterior distribution on travel-time corrections may be less accurate when the GT level of the calibration event is larger.

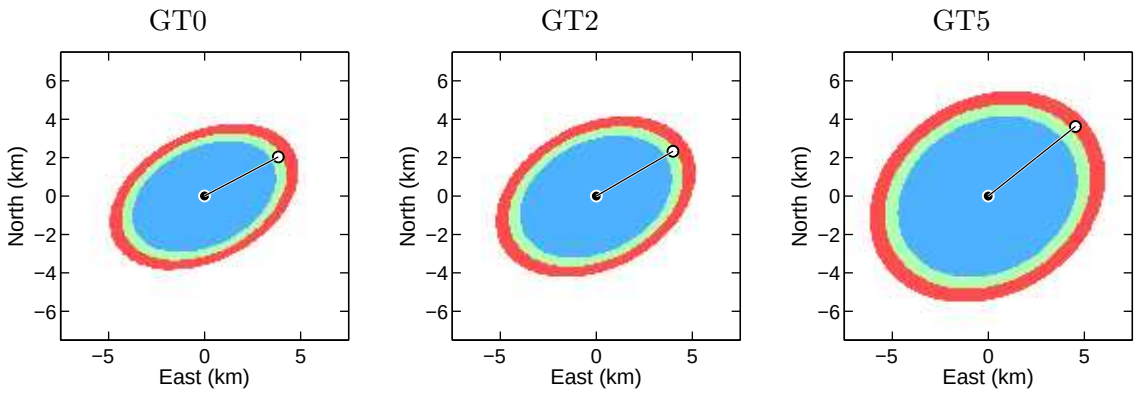

Figure 4.3: Numerical confidence regions for the Pahute Mesa target event, computed with the two-stage algorithm. The results for different GT levels assigned to the ground-truth calibration event are compared (as in Figure 3.1). Gaussian data errors were assumed.

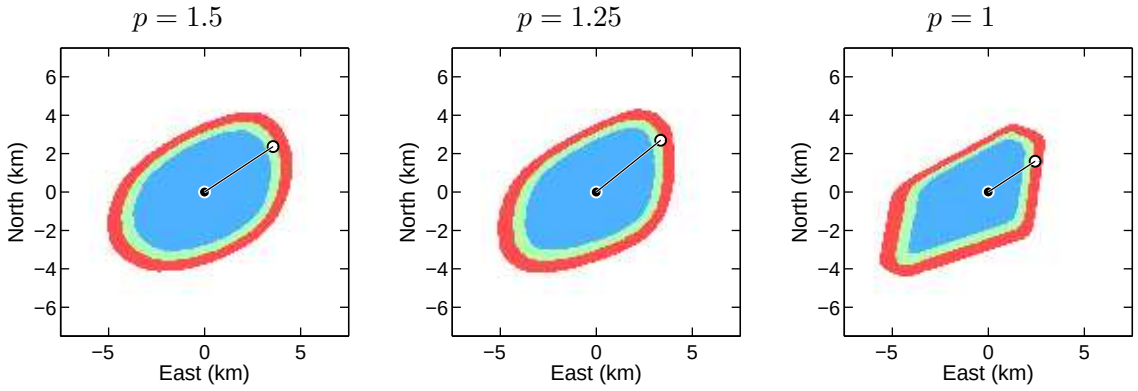

Figure 4.4: Numerical confidence regions for the Pahute Mesa target event obtained under the assumption of non-Gaussian data errors. The order of the generalized Gaussian distribution is $p = 1.5$ (left), $p = 1.25$ (center) or $p = 1$ (right). The epicenter of the ground-truth calibration event was taken to be GT0.

Chapter 5

Conclusions and Recommendations

This project developed a rigorous theoretical framework for uncertainty analysis in seismic event location, linking travel-time prediction errors (model errors) to uncertainty in travel-time calibration. A complete uncertainty analysis requires considering the joint location/calibration inverse problem, but the implementation of such an analysis, without simplifying approximations, is practical only for simple methods of calibration (e.g. the station time terms assumed in basic multiple-event location).

The assumption that calibration data determine a Gaussian distribution on travel-time corrections allows one to factor the joint location/calibration problem into two stages. The *calibration* stage finds the Gaussian distribution, while the follow-up stage of target-event *location* uses the distribution as a prior on the corrections affecting the target event. We have implemented and tested the two-stage approach for the basic multiple-event location problem and confirmed that it is significantly more efficient than the complete joint inversion approach. The confidence regions in Figures 4.3 and 4.4 required at most 1 minute of CPU time for the calibration stage and less than 5 seconds to compute each confidence region in the location stage. This indicates that a rigorous location uncertainty analysis employing complex parameterizations of travel-time corrections is a feasible task using the two-stage approach. Of particular interest is the case of tomographic corrections, parametrized by 3-D velocity models. The recent work by Rodi and Myers (2007) provides a starting point for implementing the calibration stage for this case.

Further research is needed to test the accuracy of the Gaussian approximation used in the two-stage uncertainty analysis The suitability of this approximation no doubt depends on the pick error distribution, the number of data and station geometry, the GT location constraints on calibration events, and the degree of nonlinearity in the travel-time forward problem. An investigation of this topic could include the exploration of other, more general approximations to the correction posterior, such as with copulas. Another important topic for further study is focal depth uncertainty. The formulation and algorithms developed in this project apply to the computation of confidence intervals on focal depth but our numerical tests considered only fixed-depth solutions and epicentral confidence regions. Aside from the importance of focal depth in nuclear monitoring, we can expect focal-depth uncertainty to be significantly affected by model errors and travel-time nonlinearity, and therefore take good advantage of the generality of our approach.

References

Anderson, T.W. (2003), *Introduction to Multivariate Statistical Analysis,* 3rd ed. John Wiley and Sons, Hoboken, NJ, 742 pp.

Ballard, S. (2002). Seismic event location using Levenberg-Marquardt least squares inversion, Sandia Report, SAND2002-3083.

Billings, S.D., M.S. Sambridge and B.L.N. Kennett (1994). Errors in hypocenter location: picking, model and magnitude dependence, *Bull. Seism. Soc. Am., 84,* 1978–1990.

Evernden, J.F. (1969). Precision of epicenters obtained by small numbers of world-wide stations, *Bull. Seism. Soc. Am., 59,* 1365–1398.

Flinn, E.A. (1965). Confidence regions and error determinations for seismic event location, *Rev. Geophys., 3,* 157–185.

Geiger, L. (1912). Probability method for the determination of earthquake epicenters from the arrival time only, *Bull. St. Louis Univ., 8*, 60–71.

Jordan, T.H. and K.A. Sverdrup (1981). Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seism. Soc. Am., 71*, 1105–1130.

Murphy, J R., W. Rodi, M. Johnson, D.D. Sultanov, T.J. Bennett, M.N. Toksöz, V. Ovtchinnikov, B.W. Barker, D.T. Reiter, A.C. Rosca and Y. Shchukin (2005). Calibration of International Monitoring System (IMS) stations in central and eastern Asia for improved seismic event location, *Bull. Seism. Soc. Am., 95,* 1535–1560.

Pavlis, G.L, and J.R. Booker (1983). Progressive multiple event location (PMEL), *Bull. Seism. Soc. Am.* 73, 1753–1777.

Rodi, W. (2006). Grid-search event location with non-Gaussian error models, *Phys. Earth. Planet. Int.,* in press.

Rodi, W., and R.L. Mackie (2001). Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion, *Geophysics, 66,* 174–187.

Rodi, W., E.R. Engdahl, E.A. Bergman, F. Waldhauser, G.L. Pavlis, H. Israelsson, J.W. Dewey and M.N. Toksöz (2002). A new grid-search multiple-event location algorithm and a comparison of methods, *Proceedings,* 24th Annual DoD/DOE Seismic Research Review.

Rodi, W. and S.C. Myers (2007). Modeling travel-time correlations based on sensitivity kernels and correlated velocity anomalies, *Proceedings,* 29th Annual DoD/DOE Monitoring Research Review.

Rodi, W. and M.N. Toksöz (2000). Grid-search techniques for seismic event location, *Proceedings,* 22nd Annual DoD/DOE Seismic Research Review.

Rodi, W., C.A. Schultz, G. Johannessen and S.C. Myers (2005). Grid-search location methods for ground-truth collection from local and regional seismic networks, *Final Technical Report*

to National Nuclear Security Administration, Department of Energy, Contract nos. DE-FC03-01SF22397 and W-7405-ENG-48.

Schultz, C.A., S.C. Myers, J. Hipp and C.J. Young (1998). Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. Seism. Soc. Am.,* 88, 1275–1288.

Walter, W.R., K.D. Smith, J. L. O'Boyle, T.F. Hauk, F. Ryall, S.D. Ruppert, S.C. Myers, M. Anderson, and D.A. Dodge (2003), Improving the fundamental understanding of regional seismic signal processing with a unique western United States dataset, *Proceedings*, 25th Seismic Research Review.

Wilcock, W.S.D. and D.R. Toomey (1991). Estimating hypocentral uncertainties for marine microearthquake surveys: a comparison of the generalized inverse and grid search methods, *Marine Geophysical Researches, 13*, 161–171.